# Probability (Part 2)

Grinnell College

February 19, 2025

# Warm-up

Consider rolling a six-sided dice with the following events:

$$A = \{2, 4\}, \qquad B = \{1, 4, 6\}$$

- What is $P(A)$?
- What is $P(B)$?
- What is $P(A \text{ and } B)$?
- If you know that $A$ has occured, what is the probability that $B$ also occured?
- If you know that $B$ has occured, what is the probability that $A$ has occured?

The **General Addition Rule** states that for events $A$ and $B$,

$$P(A \text{ or } B) = P(A) + P(B) - P(A \text{ and } B)$$

If $A$ and $B$ are **disjoint** then

$$P(A \text{ or } B) = P(A) + P(B)$$

# Independence

We say that two *random processes* are **independent** if the outcome of one process provides no information about the outcome of another

Examples include:

- Flipping a coin multiple times
- Rolling a red and white dice together
- Sampling different colored marbles from a jar *with* replacement

# Independence

If two random processes $A$ and $B$ two different and *independent* processes, then the probability that both $A$ and $B$ occur is

$$P(A \text{ and } B) = P(A) \times P(B).$$

This is known as the **Multiplication Rule**

# Conditional Statistics

When two random processes are *not* independent, we say that they are *associated*, meaning that the occurence of one event provides information related to another event.

For example, if we know that event $B$ has occured and we want to assess the probability that $A$ also occured, we are looking for *the probability of $A$ given $B$* denoted $P(A|B)$

If $A$ and $B$ are independent, that is, if $B$ occuring tells us nothing new about the probability of $A$, then

$$P(A|B) = P(A)$$

# Example

Suppose we took a collection of 9 individuals, giving 5 of them a placebo and 4 of them an active drug. We then recorded how many from each group were sick once in the next 100 days

| | Sick | Not Sick | Total |
|---|---|---|---|
| Placebo | 3 | 2 | 5 |
| Drug | 1 | 3 | 4 |
| Total | 4 | 5 | 9 |

| | Sick | Not Sick | Total |
|---|---|---|---|
| Placebo | 3 | 2 | 5 |
| Drug | 1 | 3 | 4 |
| Total | 4 | 5 | 9 |

From this, we may ask ourselves: what is the probability that a randomly selected person is sick?

$$P(\text{Sick}) = \frac{\#\text{ of sick people}}{\text{Total }\#\text{ of people}} = \frac{4}{9} = 0.44$$

Or we may ask: what is the probability that a randomly selected person is sick and has been given a placebo?

$$\begin{aligned} P(\text{Is sick and got placebo}) &= \frac{\#\text{ of sick people with placebo}}{\text{Total }\#\text{ of people}} \\ &= \frac{3}{9} = 0.33 \end{aligned}$$

This brings us to a trio of definitions:

The **marginal probability** of a sample describes the probability of a *single* variable without regard to others, i.e., the probability of event $A$ is $P(A)$

The **joint probability** of a sample describes the probabilities for two or more outcomes together, i.e., the probability of events $A$ and $B$ both is $P(A \text{ and } B)$

**Conditional probability** describes the probability of one event based on the assumed outcome of another, i.e., the conditional probability of event $A$ *given* $B$ is denoted $P(A|B)$

| | Sick | Not Sick | Total |
|---|---|---|---|
| Placebo | 3 | 2 | 5 |
| Drug | 1 | 3 | 4 |
| Total | 4 | 5 | 9 |

Does knowing that a person received a placebo change our estimate of the probability that they were sick

$$
\begin{aligned}
P(\text{Is sick } \textit{given} \text{ placebo}) &= P(\text{Sick} \mid \text{Placebo}) \\
&= \frac{\#\text{ of sick people with placebo}}{\text{Total } \#\text{ given placebo}} \\
&= \frac{3}{5} \\
&= 0.6
\end{aligned}
$$

However, we can also compute this same thing from *marginal* and *joint* probabilities

| | Actual Photo | | |
|---|---|---|---|
| | Sick | Not Sick | Marginal Prob |
| Placebo | 0.33 | 0.22 | 0.55 |
| Drug | 0.12 | 0.33 | 0.45 |
| Marginal Prob | 0.45 | 0.55 | 1 |

$$
\begin{aligned}
P(\text{Is sick } \textit{given} \text{ placebo}) &= P(\text{Sick} \mid \text{Placebo}) \\
&= \frac{P(\text{Sick and Placebo})}{P(\text{Placebo})} \\
&= \frac{0.33}{0.55} \\
&= 0.6
\end{aligned}
$$

# Conditional Probability

From this, we have the following relation:

$$P(A|B) = \frac{P(A \text{ and } B)}{P(B)}$$

Rewriting this gives us the **General Multipliction Rule**:

$$P(A \text{ and } B) = P(A|B) \times P(B)$$

Note that when $A$ and $B$ are independent, $P(A|B) = P(A)$, giving us our original **Multipliction Rule**:

$$P(A \text{ and } B) = P(A) \times P(B)$$

OpenIntro Statistics, 4th Edition